

Processamento de Fala 2012/13

3º Mini-teste

23 de Maio de 2013

Identifique o seu teste colocando o seu nome e número de aluno no espaço reservado na última página. Só são aceites respostas às questões de escolha múltipla assinaladas no local apropriado nessa página. As respostas erradas a perguntas de escolha múltipla serão penalizadas.

1. Diga se as seguintes afirmações são verdadeiras ou falsas:

- (a) As técnicas de *crowd sourcing* são frequentemente utilizadas para treinar modelos acústicos, pedindo-se aos *Turkers* para fazer as transcrições fonéticas de segmentos de fala.
- (b) A perplexidade de um dado modelo de língua é máxima quando todas as palavras são equiprováveis.
- (c) Em reconhecimento do orador, as características perceptivas aprendidas são mais fáceis de extrair do que as físicas.
- (d) Em modelos de Markov contínuos, adoptam-se frequentemente matrizes diagonais por forma a aumentar a velocidade do reconhecedor.
- (e) Os coeficientes cepstrais usados em reconhecimento de fala incluem informação sobre a excitação.
- (f) Os desvios da pronúncia canónica ocorrem sobretudo na rima.

2. Indique o nome dos coeficientes que representam uma filtragem temporal das trajectórias.

3. Como se chamam as figuras que mostram o desempenho de um sistema de detecção de palavras chave? quais os respectivos eixos?

4. Quais as etapas do algoritmo *binary split*?

5. Indique um método de aprendizagem automática frequentemente utilizado para modelar pronúncias alternativas.

6. Indique duas medidas de distância usadas em reconhecimento de fala, para além da distância euclidiana.

7. Considere um sistema de reconhecimento de fala contínua e vocabulário extenso, independente do orador aplicado à transcrição de notícias televisivas. Para uma dada notícia, a transcrição manual de 3 segmentos foi a seguinte:

- a) *No entanto, muitos alunos chegam ao fim do secundário sem sequer saber, afinal, o que é a Constituição.*
- b) *O tempo dado à Constituição da República nos programas de História do oitavo e nono ano, não deixou memória.*
- c) *Nunca fui grande fã de História, por isso basicamente decorava o que isso era para esquecer no dia seguinte ao teste.*
- d) *Está escrito onde?*

A transcrição automática produzida pelo reconhecedor foi:

- a) *No entanto, muitos alunos chegam ao fim do secundário. Sem sequer saber, afinal, o que é a Constituição.*
- b) *O tempo estava a Constituição da República nos programas de História do muito sabe nono ano, não deixa memória.*
- c) *Nunca fui grande fã de História, por isso curável que isso era para esquecer no dia seguinte o dez.*
- d) *E se a escrito onde.*

Ignorando a pontuação e capitalização, complete os valores de H ("correct"), D ("deletions"), S ("substitutions"), I ("insertions"), N ("total"), %Corr, %Acc e %WER correspondentes para cada um dos segmentos. Diga também, para cada um dos segmentos, se correspondem a: fala preparada em ambiente limpo (L), fala preparada em ambiente ruidoso (R), ou fala espontânea em ambiente ruidoso (E).

8. Considere o corpus de treino composto pelas seguintes frases:

O Obelix gosta de javalis. O Asterix gosta mais da poção mágica. O Panoramix prepara a poção mágica. Os romanos conquistaram a Gália. O Obelix e o Asterix não gostam dos soldados romanos.

Considere a frase de teste:

Os soldados romanos não gostam da poção.

- Calcule o número de unigramas, bigramas e trigramas do corpus de treino, e a dimensão do vocabulário.
- Calcule a probabilidade da frase de teste usando um modelo de unigramas.
- Calcule a probabilidade da frase de teste usando um modelo de bigramas sem alisamento.
- Calcule a probabilidade da frase de teste usando um modelo de bigramas com alisamento do tipo add-one.
- Existe algum quadrigrama no corpus de treino com mais que uma ocorrência? Se sim, indique-o e calcule a sua probabilidade.
- Considere que existe uma classe sintática *gostar* constituída por todas as formas verbais deste verbo. Qual o número de bigramas diferentes?
- Escreva uma frase o mais longa possível, cuja probabilidade de acordo com um modelo de bigramas seja não nula.

(Nota: Indique as fracções correspondentes a cada n-grama, sem calcular o produto final.)

9. Pretende-se determinar a temperatura média de um dado lugar na terra num passado distante em que não existiam ainda termómetros. Para simplificar, consideraremos apenas duas temperaturas: quente (Q) e frio (F). A probabilidade de um ano quente se suceder a outro ano quente é de 0.7. A probabilidade de um ano frio se suceder a outro ano frio é de 0.6. A transição entre estados Q e F será modelada por um modelo de Markov, em que o próximo estado depende apenas do estado corrente. Como não existiam termómetros, estes estados não são directamente observáveis. Todavia, existe uma relação probabilística entre o tamanho dos anéis de crescimento das árvores (pequenos (P), médios (M) ou grandes (G)) e a temperatura, ilustrada pela matriz seguinte:

	P	M	G
Q	0.1	0.4	0.5
F	0.7	0.2	0.1

Suponha ainda que a distribuição inicial de estados é $\text{prob_inicial} = \{Q: 0.6, F: 0.4\}$

- Determine qual a probabilidade da sequência de estados QQFF estar associada à sequência de observações (P, M, P, G).
- Trata-se de um modelo de Markov de primeira ou segunda ordem?

(Exemplo adaptado de M. Stamp, A Revealing Introduction to Hidden Markov Models, 2012)

10. Considere um sistema de reconhecimento de palavras isoladas com vocabulário limitado (7 notas musicais).

- Desenhe a topologia de um modelo de Markov adequado que preveja saltos entre estados consecutivos.
- Diga quantas probabilidade de transição há que treinar para cada modelo.
- Suponha que para uma dada locução de teste, o sistema deu os seguintes valores de probabilidades para os 7 modelos: 0.4, 0.3, 0.3, 0.3, 0.8, 0.3, 0.5. Calcule a medida de confiança para a quinta palavra.

Respostas

Nome:	
Número:	

1. (1,8 val.)

a	b	c	d	e	f

2 a 6 (1 / 1,5 / 2 / 1 / 1,2 valores)

2	
3	
4	
5	
6	

7. (3,0 val.)

H	D	S	I	N	% Corr	% Acc	% WER	L/R/E

8. (3,5 val.)

a)	
b)	
c)	
d)	
e)	
f)	
g)	

9. (2 val.) e 10. (3 val.) Responder na página seguinte.

